DOCUMENT RESUME

ED 456 141 TM 033 203

AUTHOR Shore, Arnold; Madaus, George; Clarke, Marguerite
TITLE Guidelines for Policy Research on Educational Testing. NBETPP Statements, Volume 1, Number 4.
INSTITUTION National Board on Educational Testing and Public Policy, Chestnut Hill, MA.
SPONS AGENCY Ford Foundation, New York, NY.
PUB DATE 2000-04-00
NOTE 9p.
AVAILABLE FROM For full text: http://www.nbetpp.bc.edu.
PUB TYPE Opinion Papers (120)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Educational Policy; *Educational Research; *Policy Formation; Public Policy; *Research Utilization

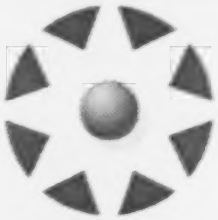
ABSTRACT

To ensure that its studies are relevant to policy making, the National Board on Educational Testing and Public Policy (NBETPP) follows certain guidelines. These are discussed in the hope that they will make the work of NBETPP more understandable and useful. The first guideline is that to be policy relevant, research must take account of the concerns of policy formulators and reformulators, those who react to or reshape the policy. Research must take into account the full range of factors affecting a policy and lay out possibilities for action. A second guidelines is that studies of testing must describe in detail what actually takes place in testing programs to ensure that policymakers and those who implement policy have common starting points. The third guideline affirms that to be policy relevant, testing studies must generate realistic policy options for incremental change and present them along with policy alternatives. The fourth guideline indicates that research on testing policy must try to forecast the implications of policy options so that those who formulate, implement, and evaluate public policy programs will appreciate the dynamics of various policy alternatives. (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 456 141

The National Board on Educational Testing and Public Policy



NBETPP

statements

Volume 1 Number 4 April 2000

# Guidelines for Policy Research on Educational Testing

Arnold Shore, George Madaus, Marguerite Clarke
National Board on Educational Testing and Public Policy
Peter S. and Carolyn A. Lynch School of Education
Boston College

Educational tests are in great demand:

- Twenty-nine states require – or soon will require – students to pass a test for graduation from high school. Twelve states have or will have tests to determine grade-to-grade promotion.
- 12.6 million children are tested in state-mandated high-stakes testing programs. An additional 4.1 million children are tested in state testing programs without high-stakes attachments.
- In light of recent court decisions and public referenda, tests are becoming more central in admissions to college and graduate school. The number of students taking a college admissions test rose to 3 million in 1999.

TM033203

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
M. Clarke
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
1

BEST COPY AVAILABLE

Despite the prevalence of educational testing, much of the technology involved is opaque to policy makers, practictioners, and the public alike. The effects of testing, positive and negative, may be more evident, but they too are often not fully understood. By conducting studies of educational testing that are accessible to lay audiences, the National Board hopes to engage all interested parties in informed debate about national, state, and local testing policy.

To ensure that our studies are relevant to policy making, we follow certain guidelines. These are discussed here in the hope that they will make our work more understandable and useful.

## Guideline I

*To be policy relevant, research must take account of the concerns of policy formulators (those who initially decide on the policy) and policy re-formulators (interested audiences who implement, react to, and/or reshape the policy). It must consider the full range of factors affecting a policy and lay out possibilities for short- and long-term action.*

Given the problems of allocating and managing scarce resources of time, money, and expertise, educational testing draws the attention of policy makers, practioners, and the public only when a crisis seems to have been reached. Currently, that point is the perceived failure of a school system that shortchanges students and the public. In testing, the current policy preference is for state-mandated tests matched to standards of attainment in an effort to hold students, teachers, schools, and districts accountable for student learning. Often the tests are part of a state-level accountability system that uses various indicators for schools and districts (e.g., attendance, graduation, and dropout rates) in order to measure "performance" and punish or reward under- or over-performing schools.

This trend is of concern to the National Board. In many instances the policies (the tests and the accountability programs that often surround them) are put in place without adequate attention to the factors that will drive the policy and how these factors can be manipulated. These so-called "actionable

**Arnold Shore** is the Executive Director of the National Board on Educational Testing and Public Policy.

**George Madaus** is a Senior Fellow with the National Board on Educational Testing and Public Policy and the Boisi Professor of Education and Public Policy in the Lynch School of Education at Boston College.

**Marguerite Clarke** is a Senior Research Associate with the National Board on Educational Testing and Public Policy.

variables" need to be identified as they are key to gaining control over the way a policy is implemented and the outcomes it produces.

One example of an actionable variable is the cut scores or achievement levels on the tests and their implications for long-term decisions about students, teachers, schools, and education systems. The points where the cut scores are set on the state examination will come to define levels of high school achievement, no matter a student's grade point average, standing in class, or attainment on other tests (e.g., Stanford 9 or the Iowa test batteries). Furthermore, they will determine in large part how a school views itself and can even affect how a community thinks about its schools and the teachers who work in them. Given the importance of cut scores, we therefore need to consider them carefully – educationally, technically, and in terms of social policy goals.

The points where the cut scores are set on the state examination will come to define levels of high school achievement, no matter a student's grade point average, standing in class, or attainment on other tests.

Other possible actionable variables in the context of high-stakes testing policy include the allocation of educational resources to schools and the types of information used to make decisions about how well a school is "performing." By conducting studies that identify the actionable variables for any policy, implementation can proceed in a more focused manner and with a greater understanding of likely short- and long-term effects.

## Guideline II

*To be policy relevant, studies of testing must describe in detail what actually takes place in testing programs, so that those involved in the making, implementation, and evaluation of policy have common starting points.*

In the policy-making process it is easy to get caught up in the rhetoric and promise of standards-based tests, exit examinations, and proficiency tests. However, to evaluate a particular testing policy properly, it is necessary to unpack what the policy actually does, determine to what extent actual outcomes match intended outcomes, and illuminate unintended outcomes, both positive and negative.

**In any policy context everything is connected as though part of a web; and an effort to formulate a policy "silver bullet" winds up spraying shot in many directions.**

Understanding the actual effects of a testing policy is critical. In any policy context everything is connected as though part of a web; and an effort to formulate a policy "silver bullet" winds up spraying shot in many directions. What actually takes place and where the effects lead are important subjects to tackle in order to separate rhetoric from reality.

For example, while teachers and superintendents are often left out of initial deliberations on standards-based reform, they are usually brought into the process at the implementation stage. In any evaluation of a testing program, regular, up-to-date information in the form of surveys of these educators (as well as community members) will be useful for understanding (1) how those who implement test policy think about tests and the decisions they must make on the basis of test outcomes; (2) how they actually use tests; and (3) what decisions they make based on test outcomes.

This type of research will help illuminate the disparities, if any, between the intended outcomes, as proposed by policy formulators, and the actual outcomes as experienced by policy implementers.

## Guideline III

*To be policy relevant, testing studies must generate realistic policy options for incremental change and present them alongside variations, or so-called policy alternatives, within options.*

In the field of policy studies perhaps no one stands taller or more influential than Charles Lindblom. Lindblom's conceptualization of the policy process in the US is seminal: all policy change is, should be, and can only be incremental in a democratic system. The notion of incremental policy change is a key starting point for testing policy analysis and a guide for providing realistic policy options developed from evidence-based research.

As we study testing programs and systems and develop policy options and alternatives within options, we need to keep in mind that variations on incremental themes are of keen

interest in a field characterized by dissention and ideological differences. A case in point is a series of studies to be undertaken by the National Board on computer-based and computer-adaptive tests. Before tests taken on the computer (computer-based) or whose content is guided interactively by real-time scoring (computer-adaptive) become commonplace, we need to examine the relationship between computerized tests and test performance, and generate options for their realistic implementation and integration into the education system. For example, studies exploring incremental policy options in the area of computerized testing need to take account of all of the following:

- **Costs** – Given current arrangements, is the test policy increasing or decreasing costs to the user or the producer? How can costs be controlled?
- **Administration** – What steps are involved in the implementation and management of the policy? Is there a necessary sequence of events or is flexibility possible at certain stages?
- **Coverage** – Whom/what will the test policy cover and when?
- **Equity** – How is the test policy affecting the range of users, especially those historically underserved by our educational systems?
- **Outcomes** – What benefits and harms is the policy producing? Are evaluation stopping-off points built into the policy implementation process so that unintended harm can be caught in time and possibly reversed or mitigated? How can benefits be maximized and harm minimized?

As we think about incremental changes and options in computer-based and computer-adaptive testing, ease of administration may be a direct tradeoff with steeply increased costs to producers (to ensure item and test security, the item bank may have to be increased many times in size); who will be covered and how is still evolving, but not systematically; and equity is a major concern, since test takers' experience with computers depends on socioeconomic level and seems to be directly related to test gains.

**As we study testing programs and systems and develop policy options and alternatives within options, we need to keep in mind that variations on incremental themes are of keen interest in a field characterized by dissention and ideological differences.**

All these questions must be addressed. The National Board and others will need to study the current effects and possibilities of computerized tests and develop incremental policy options and alternatives for the consideration of decision makers and the public.

## Guideline IV

*To be policy relevant, research on testing policy must try to forecast the implications of policy options so that those who formulate, implement, and evaluate public policy programs will appreciate the probable dynamics of various policy alternatives.*

As with every decision of consequence, time is a major concern in educational test policy: how long will it take to bring the policy on line, and where will it lead? We are often called upon to estimate the probable intended and unintended consequences as policy unfolds over time.

Projections are a mixture of evidence, experience, and intuition. It is probably impossible to avoid the introduction of some personal values or biases in such estimates. Thus, the researcher needs to make explicit from the start his or her position on testing and educational improvement. The National Board's position is this:

- Testing is a technology
- Like all technology, testing has flaws and limits
- It is important is to recognize the imperfections of testing and try to maximize its value as a source of information and minimize its harm through distortion

Let us take the case of accountability systems for states. In these systems tremendous emphasis is placed on continuous educational improvement. Yet with time, the improvement rate, whether cast as a percentage or a number of students achieving a cut-score category (e.g., "proficient" or "basic") will inevitably slow.

**As social scientists working in the area of test policy, we need to project carefully and thoughtfully what reasonable educational improvement over time might look like.**

As social scientists working in the area of test policy, we need to project carefully and thoughtfully what reasonable educational improvement over time might look like. These projections will need to look at past rates of improvement and to project growth in related areas, such as professional development of teachers, availability of resources, and the like.

In a full analysis, the projections would set certain points at which the community, political decision makers, and educators further react to the improvements achieved. In turn, these reactions would be figured into further projections. These steps would help illuminate what levels of educational progress can be sustained over what periods of time with what resources.

## Conclusion

We trust that the guidelines we follow at the National Board in our studies will help others to understand our work. We would welcome any reactions to these guidelines and to the National Board research reported in other publications in this series.

NBETPP Home
NBETPP
Visit us on the World Wide Web at nbetpp.bc.edu for more articles, the latest educational testing news, and information on NBETPP.

BEST COPY AVAILABLE

## About the National Board on Educational Testing and Public Policy

Created as an independent monitoring system for assessment in America, the National Board on Educational Testing and Public Policy is located in the Peter S. and Carolyn A. Lynch School of Education at Boston College. The National Board provides research-based test information for policy decision making, with special attention to groups historically underserved by the educational systems of our country. Specifically, the National Board

- Monitors testing programs, policies, and products
- Evaluates the benefits and costs of testing programs in operation
- Assesses the extent to which professional standards for test development and use are met in practice

This National Board publication series is supported by a grant from the Ford Foundation.

The National Board on Educational Testing and Public Policy
Lynch School of Education, Boston College
Chestnut Hill, MA 02467

Telephone: (617)552-4521 • Fax: (617)552-8419
Email: nbetpp@bc.edu

Visit our website at **nbetpp.bc.edu** for more articles, the latest educational news, and for more information about NBETPP.



NBETPP

## The Board of Directors

**Peter Lynch**
Vice Chairman
Fidelity Management and Research

**Paul LeMahieu**
Superintendent of Education
State of Hawaii

**Donald Stewart**
President and CEO
The Chicago Community Trust

**Antonia Hernandez**
President and General Council
Mexican American Legal Defense and Educational Fund

**Faith Smith**
President
Native American Educational Services

BOSTON COLLEGE